DOCUMENT RESUME

ED 408 335 TM 026 588

AUTHOR Klein, Lawrence W.; Catizone, Carmen A.
TITLE Validation of the Blueprint for an International Licensure Examination.
PUB DATE Mar 97
NOTE 23p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Chicago, IL, March 25-27, 1997).
PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Foreign Countries; International Education; *Licensing Examinations (Professions); Pharmaceutical Education; *Pharmacy; Surveys; *Test Construction; Test Items; *Validity
IDENTIFIERS *Canada; United States

ABSTRACT

The National Association of Boards of Pharmacy (NABP) manages the National Association of Boards of Pharmacy Licensure Examination (NABPLEX). A new Scope of Pharmacy Practice study was completed in 1994, and representatives of NABP and the National Association of Pharmacy Regulatory Authorities of Canada met in 1995 to discuss the possibility of creating an international pharmacy licensing examination. The first step was to determine whether a common examination blueprint between the two countries was feasible. The NABPLEX competency statements were updated based on the results of the 1994 study, and they were then reviewed independently by representatives of each country to assess their applicability to the country. A survey instrument was developed to assess the relative importance of each competency statement as it relates to entry-level pharmacy practice. Results were analyzed separately for the 1,019 American respondents, 358 Canadian respondents, and both groups combined, and a joint examination blueprint was approved by both groups. The focus of this paper is on the process used to establish the joint blueprint, rather than on the blueprint itself. An appendix contains charts of survey responses. (Contains two tables, three appendix charts, and seven references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 408 335

Validation of the Blueprint for an International Licensure Examination

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Lawrence W. Klein
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Lawrence W. Klein
Klein & Associates

Carmen A. Catizone
National Association of Boards of Pharmacy

Presented at the Annual Conference of the
National Council on Measurement in Education (NCME)

March 1997
Chicago, Illinois

TM026588

BEST COPY AVAILABLE

Validation of the Blueprint for an International Licensure Examination

Lawrence W. Klein
Klein & Associates

Carmen A. Catizone
National Association of Boards of Pharmacy

Abstract

The National Association of Boards of Pharmacy (NABP) manages the National Association of Boards of Pharmacy Licensure Examination (NABPLEX). A new Scope of Pharmacy Practice Study was completed in 1994, and representatives of NABP and the National Association of Pharmacy Regulatory Authorities (NAPRA) of Canada met in 1995 to discuss the possibility of creating an international pharmacy licensing examination. The first step was to determine whether a common examination blueprint between the two countries was feasible. The NABPLEX competency statements were updated based on the results of the 1994 study, and they were then reviewed independently by representatives of each country to assess their applicability to each country. A survey instrument was developed to assess the relative importance of each competency statement as it relates to entry-level pharmacy practice. Results were analyzed separately for American respondents, Canadian respondents, and both groups combined, and a joint examination blueprint was approved by both groups. The focus of this paper is on the process used to establish the joint blueprint, rather than on the blueprint itself.

Introduction

The NABPLEX is a licensing examination designed to protect the health and welfare of the public by assessing the competence of candidates to function safely as entry-level pharmacists. In order to allow valid inferences to be made about candidate performance on the job on the basis of their performance on the examination, however, it is critical that the examination focus on the tasks that are required of entry-level practitioners, and that the most important job activities carry the greatest weight in the examination. Much has been written, and continues to be written, about the importance of conducting empirical studies as a basis for assigning weights to the topics covered in a credentialing examination.

Standard 11.1 of the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1985), for example, specifies that:

> "The content domain to be covered by a licensure or certification test should be defined clearly and explained in terms of the importance of the content for competent performance in an occupation. A rationale should be provided to support a claim that the knowledge or skills being assessed are required for competent performance in an occupation... " (p. 64).

The comment related to that standard specifies that "job analyses provide the primary basis for defining the content domain" (p. 64). Procedures for establishing weights on the basis of job analysis data, or "practice analysis" data as it has been described more recently (Kane, 1997), have been described in a number of articles. Kane, Kingsbury, Colton, & Estes (1989), provided a thorough explanation of the importance of controlling the relative contributions of criticality and frequency ratings in establishing estimates of importance, and Lunz, Stahl, & James (1989) provided a concise description of how a Rasch rating scale analysis could be applied to translate the results of a job analysis into test specifications. The issues addressed in each of these articles were considered in depth in devising the analyses for the present study.

A study was completed in 1994 to delineate the current job or practice-related activities of pharmacists. It was sponsored jointly by the American Association of Colleges of Pharmacy (AACP), American Pharmaceutical Association (APhA), American Society of Hospital Pharmacists (ASHP), and National Association of Boards of Pharmacy (NABP). The 1994 study did not, however, focus specifically on entry-level pharmacy practice, nor did it address the "importance" of practice-related activities in a way that led to development of a test blueprint. Hence the need for the present study.

The present study grew out of the need to review the NABPLEX blueprint in light of the results of the 1994 study, and to update it if necessary. In addition, there had been discussions between representatives of the United States and Canada regarding the feasibility of developing a single licensure examination that would meet the needs of both countries for assessing the competence of entry-level pharmacists. Consistent with the goals of the North American Free Trade Agreement (NAFTA), creation of such an examination would help to remove the barriers associated with allowing qualified pharmacists to work in either country.

The purpose of the present study was first to assess the feasibility of establishing a common blueprint for an international pharmacy licensure examination that would be appropriate for use in both the United States and Canada, and if appropriate, to document the process by which such a blueprint could be established.

Methodology

As used here, the term "test blueprint" includes both the topics to be covered by an examination, and the weights assigned to each topic. As a result, the feasibility of a common blueprint was assessed in two phases. Phase 1 involved creating competency statements that content experts from both countries agreed delineated those aspects of pharmacy practice that have an impact on protecting the health and welfare of the public. Phase 2 involved assessing the relative importance of each competency statement, and determining whether the weights that would be assigned to each would be the same for both countries.

Developing Competency Statements

The existing NABPLEX competency statements were reviewed in light of the results of the 1994 study, and revised wherever necessary to ensure that they would continue to delineate contemporary entry-level pharmacy practice. The revised competencies were reviewed independently by the NABPLEX Review Committee (NRC), the Advisory Committee on Examinations (ACE), representatives of NAPRA, and NABP staff. Representatives of ACE and NAPRA then came together to resolve any discrepancies that arose from the separate deliberations, and to discuss whether the competency statements adequately delineated entry-level pharmacy practice in both countries.

Based on discussion during the meeting, it was clear that all of the competency statements applied to both countries, and neither group anticipated that additional competency statements would be required in order adequately delineate entry-level pharmacy practice. From the perspective of topics to be included in the blueprint, then, a single blueprint for both countries was seen as definitely feasible. Based on that, the second phase of the feasibility study was undertaken, determining whether a single set of weights for the competency statements would apply to both countries.

Assigning Weights to the Competency Statements

In order to gather data regarding the relative importance of the various competency and subcompetency statements that delineated entry-level pharmacy practice, the revised competencies were transformed into a survey instrument. A hierarchical structure was established by subdividing pharmacy practice into three major functional areas, and further subdividing those areas into competency statements and subcompetency statements. There were a total of nine competency statements, and thirty-nine subcompetency statements. A complete list of these is included as Appendix A. A review of the competencies within the survey instrument shows that the three major competency areas were very broad and general, the nine competency statements were more specific, but still fairly broad, and the 39

subcompetency statements were more specific yet.

An article by Kane, Kingsbury, Colton, & Estes (1989) provided practical guidelines for assessing the relative importance of practice-related activities.

> "The relative importance of any activity in practice will depend on the frequency of the activity (how often it is performed) and the criticality of the activity (the difference that it makes in terms of client outcomes). The results of the job analysis can be summarized in terms of the average frequency of occurrence of each activity over respondents and the average rating of criticality over respondents. The central task is then to combine average frequency and average criticality in order to get an overall index of the importance of the activity." (p. 19)

This is consistent with the expectation that, in order to be considered competent, a practitioner should make very few if any serious errors, and few errors of any kind that have an impact on the health and welfare of the public. From that perspective, ratings of frequency and criticality are both indicators of the overall importance of an activity.

In order to assess the relative importance of the activities delineated in the survey then, five-point rating scales were established to assess the criticality (i.e., seriousness of consequences) of each delineated activity, and to assess how frequently an entry-level pharmacist would be likely to perform the activity. In addition, the survey included a section on the demographic and professional characteristics of the respondents, and a section for qualitative comments about the survey. Samples of the rating scales and related instructions are included below:

Please use the following Criticality and Frequency rating scales in all of your ratings:

Criticality	*Frequency*
Generally, how serious are the consequences (e.g., harm to patient) if the competency is not performed properly?	*How often does an entry-level pharmacist in your practice setting perform the competency?*
1 Not serious (e.g., has no effect) *2 Minimally serious (e.g., causes inconvenience)* *3 Moderately serious (e.g., hinders therapeutic progress)* *4 Highly serious (e.g., worsens the patient's condition)* *5 Critically serious (e.g., is life threatening)*	*1 Very rarely (e.g., monthly or less)* *2 Rarely (e.g., weekly)* *3 Occasionally (e.g., daily)* *4 Often (e.g., hourly)* *5 Very often (e.g., many times per hour)*

For example, consider the competency "Evaluate drug therapy for the presence of pharmacotherapeutic

duplications and interactions." Although it is true that some drug duplications and/or interactions may be without clinical impact while others may be literally life threatening, please form your judgment by integrating the range of possibilities – if, overall, you believe that, generally, therapeutic progress will be hindered if drug duplications and/or interactions are not evaluated and identified, circle the "3" in the column labeled "Criticality." If in your practice setting, entry-level pharmacists evaluate drug therapy for the presence of duplication/interactions often (e.g. hourly), circle the "4" in the column labeled "Frequency." The table below illustrates how to mark your desired ratings.

Competency	*Criticality Rating*	*Frequency Rating*
Evaluate drug therapy for the presence of pharmacotherapeutic duplications and interactions	*1 2 (3) 4 5*	*1 2 3 (4) 5*

You may choose to read the competencies and make both your criticality and your frequency ratings at the same time, or you may choose to first rate the criticality of all of the competencies, and then go through the competencies making your frequency ratings.

After the survey instrument had been finalized it was pilot tested on a small group of licensed pharmacists to ensure that the instructions were clear and complete, and the rating scales were workable for delineating the competency areas of interest. Representatives of NAPRA produced French and English versions of the survey that were appropriate for distribution in Canada, taking into account differences between the spelling and language conventions of the two countries. Based on the results of the pilot test, no changes to the survey instrument were required. The number of French surveys completed and returned was, however, very small and insufficient to support meaningful separate analyses of the ratings.

An important feature of the survey instrument was the fact that rating scales were attached to both the competency statements and the subcompetency statements. This permitted two ways of generating weights for the competencies. Weights for the nine competency statements could be generated directly on the basis of the ratings of frequency and criticality for those statements. Alternatively, weights could be generated first for the 39 subcompetency statements based on their ratings of frequency and criticality, and those weights could then be combined by summing them appropriately to obtain weights for the nine competency statements. Because of the hierarchical nature of the survey instrument, working from the more general statements to the more specific, i.e., from the top of the hierarchical structure down, the analyses that began with ratings of competency statements are referred to as "top down" analyses, and those that began with ratings of the subcompetency statements are referred to as "bottom up" analyses. In a true top down analysis, rating scales could also have been attached to the three major competency areas. These were, however, judged to be too general to be amenable to ratings of frequency and criticality in the present study.

Distributing and Retrieving the Surveys

In the United States, surveys were distributed to a geographically representative random sample of 5,600 licensed pharmacists, which was drawn from the Pharmacy Manpower Database of 126,286 records. Thus the sample included approximately 4.4% of the licensed pharmacists in the database. In addition, the drafted competency statements were distributed to all state boards of pharmacy for comment, and a press release, accompanied by a copy of the survey, was distributed to all boards of pharmacy, schools and colleges of pharmacy, state and national pharmacy associations, and the pharmacy press.

In Canada, surveys were distributed to a geographically representative sample of licensed pharmacists, based on information provided by NAPRA. The total number of licensed pharmacists in each province or territory was estimated, and a random sample of 10% of the licensed practitioners in each area was targeted to receive the survey, with a minimum of 40 per jurisdiction wherever possible. NAPRA assumed responsibility for distributing and retrieving the surveys in Canada, and forwarded the completed surveys to NABP for processing. The total number of Canadian surveys sent was 1,738. The larger sampling percentage in Canada compared to the United States was intentional, in part to help ensure that sufficient completed surveys would be available to support the analyses based on Canadian respondents only, and in part to ensure that the Canadian responses would not be overwhelmed when they were combined with the American responses.

In order to improve response rates, a letter was sent from the Executive Directors of NABP and NAPRA in advance of the survey, to all potential participants. The purpose of the letter was to explain the importance of the survey and to encourage participation. A follow-up postcard was also sent two weeks before the surveys were due, encouraging people who had not returned the completed surveys to do so. Postage-paid return envelopes were included with the surveys, to simplify the respondents' task in returning them.

Analyzing the Completed Surveys

The purpose of the survey was to assess the relative importance of the competency and subcompetency statements. Although the testing standards provide some general guidelines regarding how to proceed toward the development of a blueprint, much remains to the discretion of the committees and researchers. Should frequency carry the same weight as criticality, for example, in assessing the relative importance of practice-related activities? Should primary consideration be given to the importance of the competency statements, or to the importance of the subcompetency statements, and does it make a difference? In the absence of clear philosophical answers to questions such as these, it was decided that the best overall approach would be to assess their empirical consequences, and take those

consequences into account in approving the final blueprint. As a result, alternative blueprints were generated on the basis of various combinations of assumptions, and presented for consideration by the panels of content experts. Their task was to decide whether one solution made more sense than the others from a practitioner perspective.

A total of 1,019 practicing pharmacists in the United States, and 358 in Canada responded to the survey. Taking into account surveys that were non-deliverable due to changes in address, and those in the samples who responded that they were not actively practicing pharmacy, this corresponded to response rates of approximately 25% in the United States, and 21% in Canada. The relatively low overall response rates were disappointing, given the steps that had been taken to encourage people to complete and return the surveys. A comparison of the demographic characteristics of the American respondents to the corresponding characteristics of the Manpower Database showed the respondents to be very representative. Similar comparisons could not be made for the Canadian sample, because a comparable Canadian database was not available. A review of the demographic characteristics of both the American and Canadian samples did, however, show that pharmacists having a wide range of experience, representing a wide range of practice settings did complete the survey. Completed surveys were received from practitioners in all fifty states and all ten provinces.

One indication that the overall results obtained were quite stable, despite the small number of Canadian respondents, was the fact that the preliminary results were generated for Canadian respondents immediately after the deadline for returning the surveys, based on 223 respondents, and again before the joint panel review meeting, based on 280 Canadian respondents. The results changed only fractionally based on the increased sample size. An additional 78 Canadian surveys were returned after the deadline for preparing materials for review by the panels of content experts, so they were not included in the preliminary analyses that were discussed during the panel review. They were later incorporated into the data set and all of the statistical analyses were rerun. The updated results were discussed with the panel via conference call, but again, incorporating the additional responses into the data set changed the results only fractionally. Adding the responses did not result in a different number of questions being assigned to any competency area, based on the assumption of a 150-item examination. Based on that, it appeared that incorporating more completed surveys into the data set would have changed the results very little if at all.

The last section of the survey provided respondents with an opportunity to recommend additional competencies, in case significant competency areas had inadvertently been omitted. If any such topics had been identified, a follow-up survey would have been conducted to establish weights for those topics. Based on a review of the comments received, however, no additional topics were required.

A Rasch rating scale analysis using the computer program *Bigsteps* (Wright, B.D., and Linacre, J.M., 1992) was run to transform the ratings of criticality and frequency associated with each of the nine competency statements onto a single linear scale. The transformed ratings were then combined as weighted averages to obtain "importance" estimates for each competency statement. As pointed out by Kane, Kingsbury, Colton, & Estes (1989),

> "The relative emphasis that should be given to frequency and criticality in determining importance is a matter of judgment. However, for licensure that is intended to protect the public from harm or unnecessary risk, criticality would seem to be of at least as much concern as frequency." (p. 20)

A number of different importance estimates were generated for each statement by varying the relative contributions of criticality and frequency to the importance variable. Specifically, ratios of criticality to frequency of 0:1, 1:1, 2:1, 4:1, and 1:0 were used to generate results. The 0:1 ratio, no weight given to criticality, and the 1:0 ratio, all weight given to criticality were considered only to describe the limiting extremes for the weights. Similarly, the 1:1 ratio, equal weight to criticality and frequency, was only considered as providing baseline data. The panel of content experts decided unanimously that, for purposes of licensure, criticality should carry more weight than frequency in assessing the importance of a competency area.

A second Rasch rating scale analysis was run on the criticality and frequency ratings for the 39 subcompetency statements, and as for the nine competency statements, the transformed ratings were then combined as weighted averages to obtain "importance" estimates for each subcompetency statement. As before, results were generated for ratios of criticality to frequency of 0:1, 1:1, 2:1, 4:1, and 1:0.

In order to transform the estimates of importance into actual numbers of questions in the examination, or percentages of the total test, the following system of inequalities was set up and several linear programming analyses were run to generate solutions to the following constraints:

- The total number of questions in the examination was to be 150.
- The number of questions or percentage of the examination assigned to a competency statement was to be greater than or equal to zero, and proportional to its estimated importance; the more important the statement, the greater the weight assigned to it.
- The number of questions or percentage of the examination assigned to a

subcompetency statement was to be greater than or equal to zero, and proportional to its estimated importance.

The linear programming analyses distributed the 150 questions across the blueprint subject to these constraints, and the results are reported in the following section as percentages of the examination. As explained in the following sections, separate analyses were run for the competency statements and the subcompetency statements.

Top Down Versus Bottom Up Delineations of Pharmacy Practice

The fact that rating scales had been attached to both competency statements and subcompetency statements in the survey made it possible to establish weights associated with pharmacy practice from two different perspectives, a "top down" view, and a "bottom up" view. To the extent that the overall results of the two approaches differed, this would provide the panel of content experts with two delineations of pharmacy practice, giving them two options for the final blueprint. One reason to think that the two approaches might produce different results was the fact that, because the competency statements tended to be more broad in scope than the subcompetency statements, they may have been more difficult to rate in terms of frequency and criticality. Another was the fact that the 39 subcompetency statements were not distributed equally across the nine competency statements. Competency statement 1.2, for example, included seven subcompetencies, whereas competency statement 3.3 included only two subcompetencies. The bottom up analyses would likely have resulted in a relatively large weight being assigned to a competency such as 1.2 simply because that weight was obtained by adding seven smaller weights. Conversely, the top down analysis may have resulted in a relatively low weight being assigned to a competency statement such as 1.2 since it did not take into account the fact that competency 1.2 covered so many subcompetencies. Because there was no single "correct" way to analyze the data, results were generated on the basis of both the top down and bottom up analyses, and presented to panels of content experts for review and consideration.

With the top down approach, pharmacy practice was divided into nine parts, based on the importance estimates associated with each of the nine competency statements. The more important the competency statement, the larger the part. Each of those parts was then subdivided into smaller parts based on the importance estimates for the subcompetency statements that comprised the particular competency statement.

With the bottom up approach, pharmacy practice was divided into 39 parts, based on the importance estimates associated with each of the thirty-nine subcompetency statements. The weights assigned to each competency statement were then obtained by summing the weights associated with the subcompetency statements that belonged to the

particular competency statement.

Complete analyses were run separately based on American respondents only, Canadian respondents only, and both groups combined. These included, for each group, running Rasch rating scale analyses to transform the ratings of criticality and frequency onto a single linear scale; generating importance estimates based on weighted averages of the transformed ratings; and generating potential examination blueprints based on top down and bottom up analyses.

Results

Results summarized in the following Tables were discussed independently with panels of content experts from the United States and Canada, and then jointly with the two groups. The purpose of the discussions was to determine whether any of the analyses produced a workable picture of entry-level pharmacy practice, and if so, whether it would be possible to adopt one that would be appropriate for delineating entry-level pharmacy practice in both countries. Results of the top down analyses are summarized in Table 1. The weights assigned to each competency statement are reported as percentages of the total test.

Table 1
Results of Top Down Analysis

	Criticality = Frequency			Criticality = 2 x Frequency			Criticality = 4 x Frequency		
	Percentage of Examination			Percentage of Examination			Percentage of Examination		
Competency Statement	US Only	Canadian Only	Combined	US Only	Canadian Only	Combined	US Only	Canadian Only	Combined
1.1	2.3	5.7	3.2	1.8	5.1	2.7	1.4	4.7	2.2
1.2	20.5	17.4	19.7	19.2	16.7	18.5	18.1	16.9	17.6
1.3	8.6	7.5	8.3	8.3	7.5	8.2	8.2	7.5	8.0
2.1	16.7	17.8	16.9	19.2	19.3	19.2	21.0	20.4	20.8
2.2	25.9	25.8	26.0	24.2	24.4	24.3	22.9	23.5	23.1
2.3	5.2	7.0	5.7	8.2	9.0	8.4	10.5	10.6	10.5
3.1	6.0	7.4	6.4	5.4	6.8	5.7	4.9	6.3	5.2
3.2	14.9	11.4	13.9	13.9	11.1	13.0	13.1	11.0	12.4
3.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Total*	100	100	100	100	100	100	100	100	100

* Totals may not add to 100.0% because of rounding.

The results of the top down analysis were not very encouraging for either group of reviewers. Neither group found any of the potential blueprints based on the top down approach to be believable representations of entry-level pharmacy practice. The most serious problem was the fact that competency area 1, which included competency statements 1.1, 1.2, and 1.3, did not receive enough weight. Even though the purpose of the survey was to establish the appropriate weights, both groups of content experts believed that, based on its scope, content area 1 would have to be the most heavily weighted if the blueprint were to accurately represent entry-level pharmacy practice.

Two findings that were noteworthy from the top down analyses were the fact that there was a good deal of similarity between the results obtained based on American respondents only and those based on Canadian respondents only, and the fact that giving additional weight to criticality compared to frequency had a small but noticeable effect on the final weights. This suggested that it was quite likely that a single blueprint could serve the needs of both countries, and that the decision regarding how much to weight criticality versus frequency ratings could be based on an assessment of the content-related implications of the various combinations.

Results based on the bottom up analyses are summarized in Table 2. As explained earlier, these results were based on the analyses of ratings associated with the subcompetency statements. The weight associated with competency statement 1.1, for example, was obtained by summing the weights of the four subcompetency statements that comprise it. Results were presented in this way in part because they would then be directly comparable to those obtained from the top down analyses and in part because the final weights at the subcompetency statement level are considered confidential. The results of the bottom up analyses showed a marked difference from those based on the top down approach.

Table 2
Results of Bottom Up Analysis

	Criticality = Frequency			Criticality = 2 x Frequency			Criticality = 4 x Frequency		
	Percentage of Examination			Percentage of Examination			Percentage of Examination		
Competency Statement	US Only	Canadian Only	Combined	US Only	Canadian Only	Combined	US Only	Canadian Only	Combined
1.1	9.2	10.3	9.4	8.8	9.8	9.0	8.5	9.4	8.7
1.2	24.9	24.5	25.2	24.2	24.7	24.6	23.6	24.6	24.0
1.3	14.1	14.8	14.4	14.6	15.9	15.0	15.0	16.9	15.5
2.1	9.1	8.1	8.7	10.5	9.3	10.2	11.7	10.5	11.4
2.2	11.2	10.9	10.8	10.5	9.5	10.0	9.9	8.3	9.3
2.3	5.4	5.2	5.1	5.9	5.1	5.6	6.4	5.1	6.0
3.1	3.9	3.6	3.6	3.7	2.9	3.3	3.5	2.3	3.0
3.2	17.5	17.4	17.8	17.0	17.6	17.4	16.6	17.6	17.0
3.3	4.9	5.2	4.9	4.9	5.2	4.9	4.9	5.2	5.0
Total*	100	100	100	100	100	100	100	100	100

* Totals may not add to 100.0% because of rounding.

Both the American and Canadian content experts believed that the results of the bottom up analyses conveyed an accurate picture of entry-level pharmacy practice. Of particular importance to them was the reversal of weight assigned to competency areas 1 and 2, compared to those based on the top down analysis. Both groups had believed very strongly that, from a content perspective, competency area 1 should have the greatest weight. That expectation was supported by the results of the bottom up analyses. Based on that, the answer was "yes" to the question of whether a suitable blueprint could be established on the basis of the survey results.

Another significant difference between the results based on the bottom up versus top down analyses was the fact that the bottom up analysis assigned some weight to competency statement 3.3, whereas the top down analysis did not. That competency statement, along with its two subcompetency statements, addressed a pharmacist's responsibility for educating patients and the public regarding wellness, disease states, and medical conditions, and content experts from both countries believed very strongly that it would be unacceptable to eliminate it from the blueprint.

BEST COPY AVAILABLE

A comparison of the weights that were obtained on the basis of American respondents only and Canadian respondents only showed only very small differences, and the weights obtained by combining both sets of responses were very representative of the separate ratings for each group. As noted for the top down analyses, increasing the contribution of criticality in comparison to frequency in estimating the importance of the competencies had a small but noticeable effect on the final weights that were obtained. The content experts from both countries agreed independently that even though the differences were small, the results based on criticality being weighted four times as heavily as frequency produced the best overall results. Based on that, the answer was again "yes" to the question of whether a single examination blueprint would meet the needs of both countries.

Working independently, and then together as a group, reviewers from both countries approved adopting the blueprint that was based on the bottom up analyses, for ratings based on both groups of respondents combined, with importance estimates based on criticality ratings being weighted four times as heavily as frequency ratings.

Summary and Discussion

This paper described the process whereby a single examination blueprint was developed and adopted as appropriate for establishing a single licensure examination to assess the competence of entry-level pharmacists in both the United States and Canada. Because the blueprint for such an examination would include both the competency statements according to which performance is to be assessed, and the weights assigned to those competency statements, the work described in this paper was completed in two phases.

Phase 1 involved determining whether a single set of competency statements would apply to both countries. If, for example, there were important competency areas that were important for practitioners in the United States, but not in Canada, or vice versa, a single blueprint would not have been viable. A set of competency statements delineating entry-level pharmacy practice was reviewed independently by panels of content experts from both countries to assess their applicability to each country. Based on a joint discussion of the results of the independent reviews, the competency statements were judged to be appropriate for use in both the United States and Canada.

As a safety precaution, the survey that was developed to establish weights for the competency statements included a section in which respondents could identify additional topics that they felt should have been included. No significant additional competency areas were identified by either American or Canadian respondents. This led support to the decision of the content experts from both countries to adopt the text of the competency statements as a basis for the examination blueprint.

Phase 2 involved determining whether assigning the same weights to the competency statements would be appropriate for assessing entry-level pharmacy practice in both the United States and Canada. A survey instrument was developed to allow pharmacy practitioners in both countries to assess the seriousness of the consequences associated with the competency and subcompetency statements, and how frequently an entry-level practitioner would likely be required to perform the activity, or address the issue. Those ratings of criticality and frequency were then combined to obtain overall estimates of the relative importance of the competency and subcompetency statements. The estimates of importance were then transformed into percentages of the total examination, with the more important statements being represented by larger percentages and less important statements represented by smaller percentages. The percentage of the examination assigned to any particular statement was proportional to its estimated importance.

Working independently, the American and Canadian panels of content experts both reached the same conclusion regarding the weights that should be adopted in establishing the blueprint. Each panel independently recommended adopting the blueprint that was based on the "bottom up" analysis, where the ratings were based on both American and Canadian respondents, and where criticality was weighted four times as heavily as frequency in estimating the importance of each competency and subcompetency statement. With the bottom up analysis, the importance of each of the thirty-nine subcompetency statements was determined and transformed into a percentage of the total examination. The percentage of the examination associated with each of the nine competency statements was then determined by summing the percentages associated with the subcompetency statements that comprised the competency statement.

Limitations of the Study

The most serious limitation relates to the low overall response rates, and particularly the response rate for Canadian practitioners. Given the steps that were taken to encourage participation, a higher response rate had been expected. A second limitation was the fact that only geographic representation was available as a stratifying variable in selecting the sample. It would have been preferable to also stratify on the basis of number of years in practice, and amount of time spent with entry-level pharmacists. This information was collected in the demographic section, but it was only available to describe the sample, not to select the sample initially. It would also have been desirable to conduct further analyses to assess whether there was any impact from having both French and English surveys available in Canada. Some preliminary analyses showed little difference between the French and English responses, but the French sample size was too small to serve as a basis for drawing reliable conclusions.

Conclusions

Overall, the procedures reported here to establish and validate a joint examination blueprint for use in two countries worked very well. Based on comments from survey respondents, and comments from the various reviewers and panel members, some of the key factors that contributed to the success of the procedures were the following:

1. Having panels from both countries review the results independently and reach their own conclusions before coming together as a group was very helpful. It enabled them to review the results objectively, without trying to accommodate the other panel, and it made it possible to establish an appropriate blueprint for one country or the other, regardless of whether both countries could agree on a single blueprint.

2. Including descriptors on the points associated with the five-point rating scales reduced the chances that respondents would apply different interpretations to the numerical values.

3. Conducting both the bottom up and top down analyses was seen as valuable even though the top down approach did not produce workable results. Given the time, effort, and expense that go into conducting a practice analysis study, it is important if at all possible to produce results that are workable. In general it seems likely that the bottom up analyses will yield more satisfactory results in studies such as these, in part because they are based on many more data points (78 in this case, versus 18 for the top down analysis), and in part because the subcompetency statements tend to be more specific, and therefore more amenable to ratings of frequency and criticality. Some reviewers indicated that the competency statements tended to be quite broad, making it more difficult to estimate their frequency and criticality. Because they were so broad, they didn't tend to be performed in isolation, but rather as part of other activities.

4. Steps taken to help improve response rates included: simplifying the survey instructions; shortening the survey; sending out advance notice of the survey to encourage participation; and sending a follow-up reminder. Even though these were not completely successful for the current study, at least every effort was being made to avoid the problems associated with low response rates.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). *Standards for educational and psychological testing.* Washington, DC: American Psychological Association.

Kane, M., (1997). Model-based practice analysis and test specifications. *Applied Measurement in Education.* Vol. 10, No. 1, pp. 5-18.

Kane, M., (1982). The validity of licensure examinations. *American Psychologist.* 37, pp. 911-918.

Kane, M., Kingsbury, C., Colton, D., & Estes, C. (1989). Combining data on criticality and frequency in developing test plans for licensure and certification examinations. *Journal of Educational Measurement.* Vol. 26, No. 1, pp. 17-27.

Lunz, M., Stahl, J., & James, K., (1989). Content validity revisited: Transforming job analysis data into test specifications. *Evaluation & the Health Professions.* Vol. 12, No. 2, pp. 193-206.

Wright, B., Linacre, M., (1992). *Bigsteps: A Rasch-model computer program.* Chicago, IL: Mesa Press.

Wright, B., & Masters, G., (1982). *Rating Scale Analysis.* Chicago, IL: Mesa Press.

Appendix A

Entry-Level Pharmacy Practice Competencies

Area 1: Manage Drug Therapy to Optimize Patient Outcomes

	Competency	Criticality Rating	Frequency Rating
1.1	**Evaluate the patient and/or patient information to determine the presence of a disease or medical condition, to determine the need for treatment and/or referral, and to identify patient-specific factors that affect health, pharmacotherapy, and/or disease management**	1 2 3 4 5	1 2 3 4 5
1.1.1	Identify and/or use instruments and techniques related to patient assessment and diagnosis	1 2 3 4 5	1 2 3 4 5
1.1.2	Identify and define the terminology, signs, and symptoms associated with diseases and medical conditions	1 2 3 4 5	1 2 3 4 5
1.1.3	Identify drug and non-drug methods of preventing and treating diseases and medical conditions	1 2 3 4 5	1 2 3 4 5
1.1.4	Identify patient factors, biosocial factors, and concurrent drug therapy that are relevant to the maintenance of wellness and the prevention or treatment of a disease or medical condition	1 2 3 4 5	1 2 3 4 5

	Competency	Criticality Rating	Frequency Rating
1.2	**Assure the appropriateness of the patient's specific pharmacotherapeutic agents, dosing regimens, dosage forms, routes of administration, and delivery systems**	Criticality Rating	Frequency Rating
1.2.1	Identify drug products by their generic, trade, and/or common names	1 2 3 4 5	1 2 3 4 5
1.2.2	Identify the known or postulated sites and mechanisms of action of pharmacotherapeutic agents	1 2 3 4 5	1 2 3 4 5
1.2.3	Evaluate drug therapy for the presence of pharmacotherapeutic duplications and interactions	1 2 3 4 5	1 2 3 4 5
1.2.4	Identify indications, contraindications, warnings, and precautions associated with a drug product's active and inactive ingredients	1 2 3 4 5	1 2 3 4 5

BEST COPY AVAILABLE

1.2.5	Identify physicochemical properties of drug substances that affect their solubility, pharmacokinetics, pharmacologic actions, and stability	1 2 3 4 5	1 2 3 4 5
1.2.6	Interpret and apply pharmacokinetic principles to calculate and determine appropriate drug dosing regimens	1 2 3 4 5	1 2 3 4 5
1.2.7	Interpret and apply biopharmaceutic principles, and the pharmaceutical characteristics of drug dosage forms and delivery systems, to assure bioavailability and enhance patient compliance	1 2 3 4 5	1 2 3 4 5

1.3	**Monitor the patient and/or patient information and manage the drug regimen to promote health and assure safe and effective pharmacotherapy**	Criticality Rating	Frequency Rating
1.3.1	Identify pharmacotherapeutic outcomes and endpoints	1 2 3 4 5	1 2 3 4 5
1.3.2	Evaluate patient information to determine the safety and effectiveness of pharmacotherapy	1 2 3 4 5	1 2 3 4 5
1.3.3	Identify, describe the mechanism of, and remedy adverse reactions and iatrogenic or drug-induced illness	1 2 3 4 5	1 2 3 4 5
1.3.4	Prevent, recognize, and remedy noncompliance and drug misuse or abuse	1 2 3 4 5	1 2 3 4 5
1.3.5	Identify and remedy interactions or contraindications with diagnostic or monitoring tests or procedures	1 2 3 4 5	1 2 3 4 5

Area 2: Assure the Safe and Accurate Preparation and Dispensing of Medications

Competency		Criticality Rating	Frequency Rating
2.1	**Perform calculations required to compound, dispense, and administer medications**	1 2 3 4 5	1 2 3 4 5
2.1.1	Calculate the quantity of medication to be compounded or dispensed; reduce and enlarge formulation quantities and calculate the quantity of ingredients needed to compound the proper amount of the preparation	1 2 3 4 5	1 2 3 4 5
2.1.2	Calculate nutritional needs and the caloric content of nutrient sources	1 2 3 4 5	1 2 3 4 5
2.1.3	Calculate the rate of drug administration	1 2 3 4 5	1 2 3 4 5
2.1.4	Calculate or convert drug concentrations, ratio strengths, and/or extent of ionization	1 2 3 4 5	1 2 3 4 5

2.2	**Select and dispense medications**	Criticality Rating	Frequency Rating
2.2.1	Determine whether a particular drug dosage strength or dosage form is commercially available, and whether it is available on a nonprescription basis	1 2 3 4 5	1 2 3 4 5
2.2.2	Identify commercially available drug products by their characteristic physical attributes	1 2 3 4 5	1 2 3 4 5
2.2.3	Identify the rationale for including excipients in the formulation of a commercial drug product, and predict the effects and adverse effects of these formulation factors	1 2 3 4 5	1 2 3 4 5
2.2.4	Interpret and apply pharmacokinetic parameters and quality assurance data to determine equivalence among manufactured drug products, and identify products for which documented evidence of inequivalence exists	1 2 3 4 5	1 2 3 4 5
2.2.5	Identify the appropriate packaging, storage, handling, and disposal of medications	1 2 3 4 5	1 2 3 4 5
2.2.6	Identify and describe the use of equipment and apparatus required to administer medications	1 2 3 4 5	1 2 3 4 5

BEST COPY AVAILABLE

2.3	**Prepare and compound extemporaneous preparations and sterile products**	Criticality Rating	Frequency Rating
2.3.1	Identify and describe techniques and procedures related to drug preparation, compounding, and quality assurance	1 2 3 4 5	1 2 3 4 5
2.3.2	Identify and use equipment necessary to prepare and extemporaneously compound medications	1 2 3 4 5	1 2 3 4 5
2.3.3	Identify the important physicochemical properties of a preparation's active and inactive ingredients; describe the mechanism of, and the characteristic evidence of incompatibility or degradation; and identify methods for achieving stabilization of the preparation	1 2 3 4 5	1 2 3 4 5

Area 3: Provide Drug Information and Promote Public Health

	Competency	Criticality Rating	Frequency Rating
3.1	**Access, evaluate, and apply information to promote optimal health care**	1 2 3 4 5	1 2 3 4 5
3.1.1	Identify the typical content and organization of specific sources of drug and health information	1 2 3 4 5	1 2 3 4 5
3.1.2	Interpret and evaluate data presented in textual, tabular, or graphic form	1 2 3 4 5	1 2 3 4 5
3.1.3	Evaluate the suitability, accuracy, and reliability of information from reference sources by explaining and evaluating the adequacy of experimental design and by applying and evaluating statistical tests and parameters	1 2 3 4 5	1 2 3 4 5

3.2	**Educate patients and health care professionals regarding prescription medications, nonprescription medications, and medical devices**	Criticality Rating	Frequency Rating
3.2.1	Provide information regarding a medication's therapeutic actions, and describe appropriate remedies to minimize the principal untoward effects resulting from drug therapy	1 2 3 4 5	1 2 3 4 5
3.2.2	Provide information regarding a medication's precautions, warnings, contraindications, and interactions with food	1 2 3 4 5	1 2 3 4 5
3.2.3	Provide information regarding the proper storage, administration, and disposal of medications	1 2 3 4 5	1 2 3 4 5
3.2.4	Identify products and describe techniques for the self-monitoring of patients' health status	1 2 3 4 5	1 2 3 4 5
3.2.5	Provide advice regarding the selection, use, and care of medical/surgical appliances or devices, durable medical equipment, and medication administration equipment	1 2 3 4 5	1 2 3 4 5

3.3	**Educate patients and the public regarding wellness, disease states, and medical conditions**	Criticality Rating	Frequency Rating
3.3.1	Provide information regarding medications used in the prevention and treatment of diseases and medical conditions, including emergency patient care	1 2 3 4 5	1 2 3 4 5
3.3.2	Provide information regarding nutrition, lifestyle, and other non-drug measures that are effective in promoting health or preventing or minimizing the progress of a disease or medical condition	1 2 3 4 5	1 2 3 4 5

BEST COPY AVAILABLE